

SKRIPSI
CLUSTERING PASIEN DIABETES DAN DIABETES
KOMPLIKASI BERDASARKAN REKAM MEDIS
ELEKTRONIK PASIEN PUSKESMAS MODOPURO
MENGGUNAKAN ALGORITMA K-MEANS



Disusun Oleh:

YOVAN FEBRIAWAN NURPRATAMA

19082010037

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
SURABAYA
2023

SKRIPSI

CLUSTERING PASIEN DIABETES DAN DIABETES KOMPLIKASI BERDASARAKAN REKAM MEDIS ELEKTRONIK PASIEN PUSKESMAS MODOPURO MENGGUNAKAN ALGORITMA K-MEANS

Disusun oleh:

Yovan Febriawan Nurpratama
19082010037

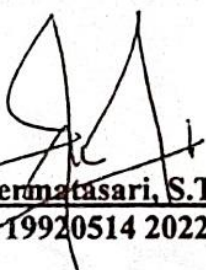
Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi
Program Studi Sistem Informasi Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Pada Tanggal 8 September 2023

Pembimbing:

1.



Dhan Satria Yudha Kartika, S.Kom., M.Kom.
NPT. 201198 60 522249

2.


Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 202203 2007

Tim Penguji

1.


Eka Dyar Wahyuni, S.Kom., M.Kom.
NIP. 19841201 2021212 005

2.


Rizka Hadiwiyanti, S.Kom., M.Kom.
NIP. 19860727 2018032 001

3.


Anita Wulansari, S.Kom., M.Kom.
NIP. 19871015 202203 2005

Mengetahui,
Dekan Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur


Prof. Dr. Ir. Navirina Hendrasari, M.T.
NIP. 19681126 199403 2 001

LEMBAR PENGESAHAN

**CLUSTERING PASIEN DIABETES DAN DIABETES KOMPLIKASI
BERDASARAKAN REKAM MEDIS ELEKTRONIK PASIEN
PUSKESMAS MODOPURO MENGGUNAKAN ALGORITMA K-MEANS**

Disusun Oleh:

Yovan Febriawan Nurpratama
19082010037

**Telah Disetujui mengikuti Ujian Negara Lisan Gelombang September
Periode 2023 pada Tanggal 8 September 2023**

Menyetujui,

Dosen Pembimbing 1



Dhian Satria Yudha Kartika, S.Kom., M.Kom.
NPT. 201198 60 522249

Dosen Pembimbing 2



Reisa Permatasari, S.T., M.Kom
NIP. 19920514 202203 2007

Mengetahui,

**Ketua Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur**



Agung Brastama Putra, S.Kom., M.Kom.
NIP. 19851124 2021211 003



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
FAKULTAS ILMU KOMPUTER**

KETERANGAN REVISI

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut:

Nama : Yovan Febriawan Nurpratama

NPM : 19082010037

Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 8 September 2023 dengan judul :

**CLUSTERING PASIEN DIABETES DAN DIABETES KOMPLIKASI
BERDASARAKAN REKAM MEDIS ELEKTRONIK PASIEN
PUSKESMAS MODOPURO MENGGUNAKAN ALGORITMA K-MEANS**

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diijinkan untuk membukukan laporan SKRIPSI dengan judul tersebut.

Surabaya, 8 September 2023

Dosen penguji yang memeriksa revisi:

1. Eka Dyar Wahyuni, S.Kom., M.Kom.
NIP. 19841201 2021212 005

{  }

2. Rizka Hadiwiyanti, S.Kom., M.Kom.
NIP. 19860727 2018032 001

{  }

3. Anita Wulansari, S.Kom., M.Kom.
NIP. 19871015 202203 2005


{  }

Mengetahui,

Dosen Pembimbing I

Dosen Pembimbing II


Dhian Satria Yudha Kartika, S.Kom., M.Kom.
NPT. 201198 60 522249


Reisa Permatasari, S.T., M.Kom
NIP. 19920514 202203 2007



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
FAKULTAS ILMU KOMPUTER**

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Yovan Febriawan Nurpratama

NPM : 19082010037

Program Studi : Sistem Informasi

Menyatakan bahwa Judul Skripsi / Tugas Akhir sebagai berikut:

**CLUSTERING PASIEN DIABETES DAN DIABETES KOMPLIKASI
BERDASARAKAN REKAM MEDIS ELEKTRONIK PASIEN
PUSKESMAS MODOPURO MENGGUNAKAN ALGORITMA K-MEANS**

Bukan merupakan plagiat dari Skripsi/Tugas Akhir / Penelitian orang lain dan juga bukan Produk / Hasil Karya yang saya beli dari orang lain.

Saya juga menyatakan bahwa Skripsi / Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di Instansi Pendidikan lain. Jika dinyatakan dikemudian hari pernyataan terbukti benar, maka Saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, 8 September 2023

Hormat Saya,



Yovan Febriawan Nurpratama

NPM. 19082010037

Judul : CLUSTERING PASIEN DIABETES DAN DIABETES
KOMPLIKASI BERDASARAKAN REKAM MEDIS
ELEKTRONIK PASIEN PUSKESMAS MODOPURO
MENGUNAKAN ALGORITMA K-MEANS

Pembimbing 1 : Dhian Satria Yudha Kartika, S.Kom., M.Kom.

Pembimbing 2 : Reisa Permatasari, S.T., M.Kom.

ABSTRAK

Puskesmas merupakan kependekan dari Pusat Kesehatan Masyarakat yang berfungsi untuk melayani kesehatan masyarakat di wilayah kelurahan dan kecamatan. Implementasi teknologi berupa Rekam Medis Elektronik (RME) memudahkan pihak puskesmas dimana sudah diterapkan pada Puskesmas Modopuro, sehingga dapat mengetahui jenis penyakit yang paling banyak diderita pasien yaitu diabetes mellitus. Pencegahan dan pencatatan dilakukan oleh pihak puskesmas untuk mencegah meningkatnya penyakit ini. Namun belum dilakukan analisa terhadap data RME untuk memperoleh informasi lebih lanjut. Oleh karena itu dilakukan *clustering* berdasarkan diagnosa penyakit menggunakan algoritma *K-Means* untuk mengetahui penyebarannya. Berkas laporan harian puskesmas yang diambil selama bulan Januari hingga April pada tahun 2023 menghasilkan *dataset* yang memiliki 627 baris. Implementasi metode *K-Means* dan teknik *One-Hot Encoding* pada *dataset* menghasilkan lima klaster. Klaster 0 sebanyak 131 orang, klaster 1 sebanyak 128 orang, klaster 2 sebanyak 109 orang, klaster 3 sebanyak 68 orang, dan klaster 4 sebanyak 191 orang. Pengujian model menggunakan metode *elbow* menghasilkan nilai $WCSS=532.6715595367782$ dan hasil evaluasi model menggunakan metode *Silhouette Coefficient* termasuk pada kategori struktur lemah dengan nilai sebesar 0.37322764931644387.

Kata Kunci:

Clustering, data mining, algoritma k-means, rekam medis elektronik, diabetes mellitus

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT, Tuhan Yang Maha Esa, atas rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan Skripsi sebagai salah satu syarat untuk menyelesaikan Program Program Studi S1/Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Dalam penyusunan Skripsi ini tidak terlepas adanya keterlibatan serta bantuan berbagai pihak baik secara moril maupun materiil, Untuk itu pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Kedua Orang Tua yang telah memberikan kasih sayang dalam mendidik, dukungan moril dan materiil, serta doa restu dalam setiap langkah penulis;
2. Prof. Dr. Ir. Akhmad Fauzi, M.MT., IPU., selaku Rektor UPN “Veteran” Jawa Timur;
3. Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer UPN “Veteran” Jawa Timur;
4. Agung Bramasta Putra, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Fakultas Ilmu Komputer UPN “Veteran” Jawa Timur;
5. Dhian Satria Yudha Kartika, S.Kom., M.Kom., selaku Dosen Pembimbing I yang telah memberikan bimbingan dan arahan dalam penyelesaian skripsi ini;
6. Reisa Permatasari S.T., M.Kom., selaku Dosen Pembimbing II yang telah memberikan bimbingan dan arahan dalam penyelesaian skripsi ini;

7. Seluruh dosen dan staf pendidikan UPN “Veteran” Jawa Timur yang telah memberikan ilmu pengetahuan dan membantu selama proses perkuliahan.
8. Bapak Drs. Nugroho Budi Sulistiya, M.Si., selaku Kepala Badan Kesatuan Bangsa dan Politik Kabupaten Mojokerto;
9. Bapak dr. Ulum Rokhmat Rokhmawan, M.H., selaku Kepala Dinas Kesehatan Kabupaten Mojokerto;
10. Bapak Edi Gandariyanto, S.Kep.Ns selaku Kepala Puskesmas Modopuro Kecamatan Mojosari;
11. Seluruh dokter dan staf kesehatan Puskesmas Modopuro;
12. Serta pihak-pihak lain yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa Skripsi ini masih memiliki beragam keterbatasan dan kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk perbaikan pada penelitian selanjutnya. Penulis berharap semoga Skripsi ini dapat memberikan kontribusi dalam menunjang perkembangan ilmu pengetahuan, khususnya ilmu komputer.

Surabaya, 10 Juli 2023

Penulis

DAFTAR ISI

LEMBAR PERNYATAAN	ii
LEMBAR PENGESAHAN	iii
KETERANGAN REVISI.....	iv
SURAT PERNYATAAN.....	v
ABSTRAK.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Dasar Teori	6
2.1.1 Data Mining.....	6
2.1.2 Rekam Medis Elektronik (RME).....	6
2.1.3 Clustering	7
2.1.4 K-Means	8
2.1.5 Python.....	8
2.1.6 Hypertext Markup Language (HTML).....	9
2.1.7 One-Hot Encoding.....	9
2.1.8 Principal Component Analysis (PCA).....	9
2.1.9 Diabetes Mellitus (DM).....	10
2.1.10 Folium.....	12
2.1.11 Metode Elbow	12
2.1.12 Silhouette Coefficient.....	13
2.2 Penelitian Terdahulu.....	14
2.2.1 Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means Clustering di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo	14
2.2.2 Cluster Analysis of Facebook Ads User for Digital Marketing Using K-Means Algorithm.....	15
2.2.3 Analisis Clustering Tingkat Keparahan Penyakit Pasien Menggunakan Algoritma K-Means (Studi Kasus di Puskesmas Bandar Seikijang)....	15
2.2.4 Visualisasi Heat Map Data Kecelakaan di Kota Palembang.....	16
BAB III METODOLOGI PENELITIAN.....	17

3.1	Studi Literatur.....	17
3.2	Analisis Kebutuhan	17
3.2.1	Kebutuhan Software dan Hardware.....	17
3.2.2	Kebutuhan Data	18
3.3	Pengumpulan Data.....	18
3.3.1	Data Primer.....	18
3.3.2	Observasi	18
3.3.3	Wawancara	19
3.4	Preprocessing Data	19
3.5	Clustering Data	19
3.5.1	Persiapan.....	20
3.5.2	Eksplorasi Dataset	20
3.5.3	Teknik One-Hot Encoding.....	21
3.5.4	Visualisasi Kemungkinan Adanya Kluster	21
3.5.5	Menentukan Nilai K Optimal	21
3.5.6	Validasi Kluster	22
3.5.7	Assign Hasil Kluster	22
3.5.8	Visualisasi Hasil Kluster.....	24
3.6	Visualisasi.....	24
BAB IV HASIL DAN PEMBAHASAN		26
4.1	Studi Literatur.....	26
4.2	Analisis Kebutuhan	26
4.2.1	Kebutuhan Software dan Hardware.....	26
4.2.2	Kebutuhan Data	27
4.3	Pengumpulan Data.....	27
4.3.1	Data Primer.....	27
4.3.2	Observasi	28
4.3.3	Wawancara	28
4.4	Preprocessing Data	29
4.4.1	Pemilihan Atribut Data.....	30
4.4.2	Ekstraksi Data.....	30
4.4.3	Data Cleansing.....	33
4.5	Clustering Data	34
4.5.1	Persiapan.....	34
4.5.2	Eksplorasi Dataset	35
4.5.3	Teknik One-Hot Encoding.....	38
4.5.4	Visualisasi Kemungkinan Adanya Kluster	39
4.5.5	Menentukan Jumlah K Paling Optimal	43
4.5.6	Validasi Kluster	46
4.5.7	Assign Hasil Kluster	47
4.5.8	Visualisasi Hasil Kluster.....	53
4.6	Visualisasi.....	61

4.6.1	Visualisasi Berdasarkan Alamat.....	61
4.6.2	Visualisasi Berdasarkan Jenis Kelamin.....	71
4.6.3	Visualisasi Berdasarkan Umur	74
4.6.4	Tampilan Website.....	78
BAB V PENUTUP.....		85
5.1	Kesimpulan.....	85
5.2	Saran	86
DAFTAR PUSTAKA		87

DAFTAR TABEL

Tabel 3.1 Kriteria Penilaian Silhouette Coefficient	22
Tabel 4.1 Daftar Pertanyaan kepada Narasumber.....	28
Tabel 4.2 Kolom yang Digunakan pada File Laporan Harian	30
Tabel 4.3 Hasil Konversi Menggunakan Teknik One-Hot Encoding dan Secara Manual.....	39
Tabel 4.4 Nilai WCSS atau SSE dalam Rentang Nilai k.....	46
Tabel 4.5 Data Acak untuk Centroid Awal.....	50
Tabel 4.6 Hasil Perhitungan Euclidean Distance.....	52
Tabel 4.7 Mengalokasikan Data ke Setiap Kluster Terdekat	52
Tabel 4.8 Centroid Baru pada Setiap Pusat Kluster	53
Tabel 4.9 Hasil Klasterisasi Setelah Proses Decode dan Keterangan Berdasarkan Kode ICD-10.....	55
Tabel 4.10 Titik Centroid Setiap Kluster pada Variabel Umur dan Jenis Kelamin	59
Tabel 4.11 Persebaran Alamat Berdasarkan Kluster.....	70
Tabel 4.12 Persebaran Jenis Kelamin Berdasarkan Kluster.....	73
Tabel 4.13 Kategori Umur Berdasarkan Depkes RI tahun 2009	74
Tabel 4.14 Persebaran Kategori Umur Berdasarkan Kluster	78

DAFTAR GAMBAR

Gambar 3.1 Alur Metodologi Penelitian.....	17
Gambar 3.2 Alur Rancangan Model Klasterisasi.....	20
Gambar 4.1 Pengumpulan Data Primer yang Berasal dari Puskesmas.....	27
Gambar 4.2 Ekstraksi Data pada File Puskesmas Menjadi File Baru dengan Format CSV	31
Gambar 4.3 Mengambil Titik Latitude dan Longitude Berdasarkan Alamat pada Tingkat Kelurahan.....	32
Gambar 4.4 Hasil Konversi Path ke Titik Latitude dan Longitude yang Dimasukkan dalam File Data DM Jan-Apr Code ICD	33
Gambar 4.5 Penanganan Data yang Tidak Memiliki Nilai atau Null	33
Gambar 4.6 Source Code Tahap Persiapan.....	35
Gambar 4.7 Source Code Eksplorasi Data.....	35
Gambar 4.8 Hasil Source Code Method Head()	36
Gambar 4.9 Hasil Source Code Method Shape.....	36
Gambar 4.10 Hasil Source Code Method Dtypes	36
Gambar 4.11 Hasil Source Code Method Describe	37
Gambar 4.12 Hasil Source Code Method Isnull.Sum	37
Gambar 4.13 Source Code Tahap Encoding	38
Gambar 4.14 Reduksi Dimensi Menggunakan Teknik PCA	40
Gambar 4.15 Hasil Source Code Teknik PCA.....	41
Gambar 4.16 Source Code Tahap Visualisasi Kemungkinan Adanya Kluster	41
Gambar 4.17 Hasil Source Code Tahap Visualisasi Kemungkinan Adanya Kluster	42
Gambar 4.18 Source Code Tahap Menentukan Jumlah K Paling Optimal	43
Gambar 4.19 Grafik Metode Elbow	45
Gambar 4.20 Source Code Tahap Validasi Kluster	46
Gambar 4.21 Hasil Source Code Tahap Validasi Kluster	47
Gambar 4.22 Source Code Tahap Assign Hasil Kluster	48
Gambar 4.23 Hasil Source Code Tahap Assign Hasil Kluster.....	48
Gambar 4.24 Source Code Mencari Nilai Centroid Setiap Kluster	49
Gambar 4.25 Hasil Source Code Mencari Nilai Centroid Setiap Kluster	50
Gambar 4.26 Source Code Decode Kolom Diagnosa.....	54
Gambar 4.27 Source Code Mengecek Diagnosa Setiap Kluster	55
Gambar 4.28 Source Code Tahap Visualisasi Hasil Kluster.....	58
Gambar 4.29 Hasil Source Code Tahap Visualisasi Hasil Kluster	60

Gambar 4.30 Source Code Visualisasi Persebaran Alamat Berbentuk HeatMap .	62
Gambar 4.31 Hasil Visualisasi Persebaran Alamat Berbentuk Peta Heatmap	64
Gambar 4.32 Visualisasi Peta Heatmap Persebaran Klaster 0	65
Gambar 4.33 Tampilan Peta Saat Ukurannya Diperbesar (Zoom In)	65
Gambar 4.34 Tampilan Peta Saat Ukurannya Diperkecil (Zoom Out)	66
Gambar 4.35 Visualisasi Peta Heatmap Persebaran Klaster 1	66
Gambar 4.36 Visualisasi Peta Heatmap Persebaran Klaster 2 Berdasarkan Alamat	67
Gambar 4.37 Visualisasi Peta Heatmap Persebaran Klaster 3	68
Gambar 4.38 Visualisasi Peta Heatmap Persebaran Klaster 4	68
Gambar 4.39 Source Code Grafik Batang Persebaran Klaster Berdasarkan Alamat	69
Gambar 4.40 Grafik Batang Persebaran Klaster Berdasarkan Alamat	70
Gambar 4.41 Source Code Grafik Batang Persebaran Klaster Berdasarkan Jenis Kelamin	72
Gambar 4.42 Grafik Batang Persebaran Klaster Berdasarkan Jenis Kelamin	73
Gambar 4.43 Source Code Pengelompokan Umur Berdasarkan Peraturan Depkes RI tahun 2009	75
Gambar 4.44 Source Code Grafik Batang Persebaran Klaster Berdasarkan Kelompok Umur	76
Gambar 4.45 Grafik Batang Persebaran Klaster Berdasarkan Kelompok Umur ..	77
Gambar 4.46 Tampilan Halaman Dashboard	79
Gambar 4.47 Tampilan Footer Pada Setiap Halaman	79
Gambar 4.48 Tampilan Halaman Persebaran Alamat Bagian Peta Heatmap	80
Gambar 4.49 Tampilan Halaman Persebaran Alamat Bagian Keterangan Peta Heatmap	80
Gambar 4.50 Tampilan Halaman Persebaran Alamat Bagian Grafik Persebaran Alamat Berdasarkan Klaster	81
Gambar 4.51 Tampilan Halaman Persebaran Alamat Bagian Grafik Persebaran Pasien Berdasarkan Kelurahan	82
Gambar 4.52 Tampilan Halaman Persebaran Alamat Bagian Keterangan Klaster	82
Gambar 4.53 Tampilan Halaman Persebaran Jenis Kelamin	83
Gambar 4.54 Tampilan Halaman Persebaran Umur	84

DAFTAR LAMPIRAN

Lampiran 1 : Surat Izin Penelitian	90
Lampiran 2 : Surat Rekomendasi Penelitian dari Badan Kesatuan Bangsa dan Politik Mojokerto	91
Lampiran 3 : Surat Rekomendasi Penelitian dari Dinas Kesehatan Mojokerto....	92
Lampiran 4 : Surat Izin Penelitian dari UPT Puskesmas Modopuro	93
Lampiran 5 : Surat Keterangan Penggunaan RME Puskesmas Modopuro.....	94
Lampiran 6 : Hasil Encoding Menggunakan Teknik One-Hot Encoding.....	95